

Application No.: 09/833,846

2

Docket No.: 08214/1200332-US2

AMENDMENTS TO THE CLAIMS

1. (Currently Amended) A method for extracting data from a network by a server, comprising:

(a) creating a database-structured query at the server based, in part, on a user input;

(b) determining a web domain address on the network from which to extract the data, ~~the web domain address having content~~; and

(c) extracting the data from a non-database structured arrangement of data directly based on the determined web domain address ~~based and~~ on the database-structured query, wherein the database-structured query is performed upon ~~at least a~~ the non-database structured arrangement of ~~content at~~ data located with the determined web domain address.

2. (Original) The method of Claim 1, wherein creating the database-structured query, further comprises, including a network address within the database-structured query indicating a starting point.

3. (Currently Amended) The method of Claim 2, wherein the determined web domain address, includes at least one universal resource locator (URL).

4. (Currently Amended) The method of Claim 2, wherein directly extracting data from the determined web domain address, further comprises, following links contained within the web domain until the links have been exhausted or following the links until a predetermined limit is reached.

5. (Original) The method of Claim 1, wherein creating the database-structured query, further comprises, creating a regular expression within the database-structured query used to determine the data to extract.

{S:\08214\1200332-US2\80036974.DOC [REDACTED]}

Application No.: 09/833,846

3

Docket No.: 08214/1200332-US2

6. (Currently Amended) The method of Claim 5, wherein directly extracting data from the determined web domain address based on the database-structured query, further comprises, matching a plurality of patterns contained within the regular expression to the ~~content~~ data to determine the data to extract.

7. (Currently Amended) The method of Claim 1, wherein creating the database-structured query, further comprises, creating a conditional expression within the database-structured query describing how to scan the ~~content~~ data at the determined web domain address for the data to extract.

8. (Currently Amended) The method of Claim 1, wherein ~~[[the]]~~ directly extracting the data from the determined web domain, further comprises:


- (a) retrieving ~~content~~ data from the determined web domain address;
- (b) reducing the retrieved ~~content~~ data to a region of interest; and
- (c) searching the region of interest for the data matching a predetermined regular expression.

9. (Currently Amended) The method of Claim 8, wherein directly extracting the data from the determined web domain, further comprises, storing the data matching the predetermined regular expression.

10. (Currently Amended) The method of Claim ~~[[9]]~~ 1, ~~wherein extracting the data from the determined web domain,~~ further comprising ~~comprises~~, reshaping the stored at least a portion of the extracted data by arranging the stored data for use by at least one data analysis software program.

11. (Currently Amended) A computer-readable medium having computer-executable instructions for extracting data from a network, the computer-executable instructions enabling actions comprising:

- (a) creating a database-structured query including a web domain address used for locating data ~~content~~, based, in part, on a user input;

{S:\08214\1200332-US2\80036974.DOC  }

Application No.: 09/833,846

4

Docket No.: 08214/1200332-US2

(b) locating the data content based on the web domain address, wherein at least a portion of the data is located [[at]] based on the web domain address provided by the database-structured query; and

(c) extracting at least the portion of the data from a non-database structured arrangement of located data directly based on the database-structured query from the located data content, wherein the database-structured query is performed upon ~~at least a~~ the non-database structured arrangement of located data content at the web domain address.

12. (Original) The computer-readable medium of Claim 11, wherein the database-structured query, further comprises, a network address included within the database-structured query indicating a starting point.

13. (Original) The computer-readable medium of Claim 12, wherein the network address, further comprises at least one universal resource locator (URL).

14. (Currently Amended) The computer-readable medium of Claim 11, wherein the web domain address, further comprises at least one link to another web domain address for locating data to extract, ~~links contained within the web domain to be followed until the links have been exhausted or until a predetermined limit is reached.~~

15. (Original) The computer-readable medium of Claim 11, wherein the database-structured query, further comprises, a regular expression within the database-structured query used to determine the data to extract.

16. (Currently Amended) The computer-readable medium of Claim 15, wherein the regular expression within the database-structured query, further comprises at least one pattern, ~~a plurality of patterns~~ used to determine the data to extract ~~from the web domain address having content~~.

17. (Currently Amended) A system for extracting data from a network comprising:

{S:\08214\1200332-US2\80036974.DOC (XXXXXXXXXXXXXXXXXXXX) }

Application No.: 09/833,846

5

Docket No.: 08214/1200332-US2

(a) a client computer system having a client network connection to the network and communicating with a server computer system, the client creating a database-structured query, based, in part, on a user input;

(b) the server computer system having a server network connection to the network and communicating with the client computer system, the server computer system further configured to perform actions, comprising:

receiving the database-structured query from the client computer system;

determining a web domain address on the network from which to extract at least a portion of the data relevant to the query, wherein the determined web domain address is provided by the database-structured query; and

extracting at least the portion of the data from a non-database structured arrangement of data directly based on the determined web domain address and the database-structured query, wherein the database-structured query is performed upon the non-database structured arrangement of data located with the determined web domain address.

~~determining a web domain address from which to extract the data based on the database-structured query, wherein at least a portion of the data is located at the web domain address; and~~

~~(c) the server computer system extracting data based on the database-structured query from the determined web domain, wherein the database-structured query is performable in an absence of a structured database arrangement of the data at the web domain address.~~

18. (Original) The system of Claim 17, wherein the database-structured query, further comprises, a network address within the database-structured query indicating a starting point.

19. (Original) The system of Claim 18, wherein the database-structured query, further comprises, a regular expression within the database-structured query used to determine the data to extract.

{S:\08214\1200332-US2\80036974.DOC ~~UNPUBLISHED~~ }

Application No.: 09/833,846

6

Docket No.: 08214/1200332-US2

20. (Currently Amended) The system of Claim 19, wherein the regular expression within the database-structured query, further comprises, ~~a plurality of patterns~~ at least one pattern used to determine the data to extract ~~from the web domain address having content.~~

21. (Original) The system of Claim 17, further comprising an editor for creating a template of regular expressions used to extract the data.

22. (Original) The system of Claim 17, further comprising at least one data extraction engine to extract the data.

23. (Original) The system of Claim 22, wherein the data extraction engine is a web crawler.

24. (Currently Amended) The method of claim 1, wherein the web domain address ~~having content~~ further comprises at least one link address ~~having~~ for locating at least a portion of the data content.

25. (Currently Amended) The computer-readable medium of claim 11, wherein the web domain address further comprises at least one link address that is followed to locate data to extract until a predetermined number of links is reached, ~~wherein at least another portion of the data is located at the at least one link address.~~

26. (Currently Amended) The system of claim 17, wherein the web domain address further comprises a link address, wherein at least another portion of the data is located ~~[[at]]~~ with the link address.

27. (Currently Amended) A method of extracting data from a network by a server, comprising:

(a) creating a database-structured query at the server based, in part, on a user input;

{S:\08214\1200332-US2\80036974.DOC [REDACTED] }

Application No.: 09/833,846

7

Docket No.: 08214/1200332-US2

(b) determining a website to search based in part on the database-structured query; and

(c) extracting at least a portion of the data relevant to the query at the web site directly based on the database-structured query, wherein the website is processed as a searchable database and the database-structured query is performed upon at least a non-database arrangement of the data at the website and the extraction of at least the portion of the data is from the non-database arrangement of the data.

28. (Currently Amended) The method of claim 27, wherein determining the website to search further comprises parsing the database-structured query to determine at least one link ~~a number of links~~ to search at the website.

29. (Currently Amended) The method of claim 27, further comprising:

(a) determining at least one other website to search based in part on the database-structured query and the provided web domain address; and

(b) extracting at least another portion of the data at the at least one other website based on the database-structured query and the provided web domain address, wherein the at least one other website includes a non-database structured arrangement of data that is processed as a searchable database.

30. (Currently Amended) The method of claim 27, wherein determining the website to search further comprises determining what data to extract based in part on the database-structured query and the provided web domain address.

31. (Previously presented) The method of claim 27, wherein extracting at least a portion of the data further comprises extracting data based in part on at least one of an HTML table, a binary file, and a matching pattern.

32. (Previously presented) The method of claim 27, further comprising, reshaping the extracted data for at least one of a database, a spreadsheet, eXtensible Markup Language (XML) display, and a statistical tool.

{S:\08214\1200332-US2\80036974.DOC 00000000000000000000000000000000 }

Application No.: 09/833,846

8

Docket No.: 08214/1200332-US2

33. (Previously presented) The method of claim 27, wherein the website is a starting website based in part on the database-structured query.

34. (Currently Amended) A method of extracting data ~~from content~~ within at least one webpage, comprising:

- (a) generating a database-structured query based, in part, on user input;
- (b) determining at least one webpage with the data content, wherein the determination of the webpage is provided by the database-structured query;
- (c) parsing the data at content of the at least one webpage in search of data that satisfies a query condition, wherein the data at the at least one webpage content is directly processed as though it is a searchable database and the query is performed upon ~~at least~~ a non-database structured arrangement of the data content at the at least one webpage;
- (d) extracting at least a portion of the data from the parsed data content of the at least one webpage that satisfies the query condition, wherein the extraction of at least the portion of the data is from the non-database structured arrangement of the data; and
- (e) reshaping the extracted data to a predetermined format.

35. (Previously presented) The method of Claim 34, wherein the search of data is performed on at least a second webpage.

36. (Currently Amended) The method of Claim 34, wherein parsing the data content of the at least one webpage further comprises following links included on the webpage and further parsing the data content of webpages determined by the links included on the webpage.

37. (Currently Amended) The method of Claim 34, wherein the structured query is generated to parse a limited portion of the data content of the at least one webpage with the limits predetermined by the user.

38. (Previously presented) The method of Claim 34, wherein the structured query is generated to search for at least one of a text string, a table, and a predefined list of words.

39. (Canceled)

{S:\08214\1200332-US2\80036974.DOC (XXXXXXXXXXXXXXXXXXXXXX) }

Application No.: 09/833,846

9

Docket No.: 08214/1200332-US2

40. (Canceled)

{S:\08214\1200332-US2\80036974.DOC 11/03/2005 18:14:20}